

Gilpatric, S. M., Vossler, C. A., & McKee, M. (2011). Regulatory enforcement with competitive endogenous audit mechanisms. *RAND Journal Of Economics (Wiley-Blackwell)*, 42(2), 292-312.  
doi:10.1111/j.1756-2171.2011.00134.x

## **Regulatory enforcement with competitive endogenous audit mechanisms**

Scott M. Gilpatric, Christian A. Vossler, and **Michael McKee**

### **Abstract**

This article adds to the regulatory compliance literature through the theoretical development and experimental testing of two endogenous audit mechanisms that use contemporaneous relative comparisons, based on disclosed information or imperfect signals of compliance effort, to generate a compliance competition among agents. This type of audit mechanism has some advantages over the more widely studied dynamic audit mechanisms that condition an agent's audit probability on past compliance, and provides an alternative explanation for the stylized fact, in many settings, that most agents are compliant most of the time even though audit rates and expected penalties are low.

## 1. Introduction

Amid the political realities of limited enforcement budgets and caps on penalties, a fundamental issue for regulators is how best to apply scarce audit resources to maximize compliance. As a collateral result, the agency wishes to productively select for audit the economic agents that are most noncompliant. This suggests employing an endogenous, rather than a random, audit mechanism. The related economics literature, the recent focus of which is on pollution regulation and tax compliance, has advocated the use of dynamic enforcement mechanisms that use information obtained through an audit to assign the agent's probability of back or future audits (Landsberger and Meilijson, 1982; Rickard, Russell, and Howroyd, 1982; Greenberg, 1984; Harrington, 1988; Harford, 1991; Alm, Cronshaw, and McKee, 1993; Raymond, 1999; Friesen, 2003; Stafford, 2008; Liu and Neilson, 2009). This literature has taken the view that static audit mechanisms are intrinsically less effective. For example, Harrington (1988) writes, "In a static analysis there is no way for the agency and the firm to react to each other's actions." Further, this literature asserts that dynamic audit models provide a better explanation for the compliance "puzzle," the stylized fact that actual compliance rates are high even though audit rates and expected penalties are low (see Alm and McKee, 1998; Harrington, 1988). Our article takes an alternative position, namely that (static) competitive endogenous audit mechanisms based on *relative* evaluation generate strong incentives for compliance and may provide a better explanation for the compliance puzzle.

Many regulations require agents to disclose information, where such disclosure will ultimately lead to increased costs, such as the levels of activities that are subject to taxation, or report information that may cause a negative market reaction. In health and safety regulation, firms are required to report work-related illnesses and injuries. In tax regulation, earners report their income through tax returns. Other regulations require agents to take costly actions, and the regulator can obtain a (imperfect) signal on compliance effort. For example, in the context of pollution regulation, there are data from air and water quality monitors. In the context of tax compliance, an individual's tax return can be compared to those within a peer group. In this article, noting that the regulator often has information from which to contemporaneously compare the relative compliance of agents, we model two audit mechanisms that implement a competition among agents to achieve higher levels of predicted compliance than random audits. The implications of the theory are tested using controlled laboratory experiments.

We frame our theory in the context of an environmental regulation setting where firms are required to disclose information on pollution, such as the case with the Toxics Release Inventory. However, these mechanisms apply more broadly, and we discuss the circumstances where they could apply to enforcement of regulations requiring actions such as pollution abatement. In contrast to many articles in the related literature, compliance effort (level of disclosure, abatement, etc.) is a continuous choice variable. The other key characteristics of our enforcement models are that audits are imperfect and that the regulator can, imperfectly, compare behavior among a group of peer firms and select for audit those firms that appear most likely to be noncompliant.

Our model assumes only that the regulator is able to make the probability a firm is audited depend to some degree on its compliance effort relative to its peers. Also, similar to Evans, Gilpatric, and Liu (2009), we model audits as imperfect, which implies that firm effort reduces at the margin the expected penalty conditional on being audited. Effort is also motivated through competition among firms to avoid being selected for an audit. This gives rise to a game that can be characterized as either a form of tournament (Lazear and Rozen, 1981) or relative performance evaluation. We show that this can yield economically significantly greater compliance effort than with purely random audits, and we discuss some significant advantages this type of endogenous auditing has relative to auditing based on compliance history. Induced competition of the sort we model may be an important reason for high compliance in the presence of a low audit frequency and small fines for noncompliance. In the equilibrium of the game, high levels of compliance are predicted even in some circumstances where purely random audits would yield zero compliance effort. These circumstances are expanded if being audited is itself a costly event for a firm.

Our tournament audit mechanism represents an interesting twist on standard tournament models, which have been used extensively to study behavior in the labor market, sporting contests, and the practice of law. In our application to compliance, the level of disclosure affects not only who gets audited but compliance costs conditional upon being audited (e.g., penalties for misreporting). This differs from the standard model where the level of effort only affects who wins but not the amount of the prize upon winning. The model developed here may be adapted to study behavior in other settings where payoffs are rank dependent but prizes are not fixed. This is generally true when agents compete to win an opportunity with a value that depends on the quality of the winning entry. Examples include advertising agencies competing to win a firm's contract and many types of research tournaments. Even a singing contest where the winner receives a recording contract appears to fit this description.

Alm, Cronshaw, and McKee (1993) and Cason and Gangadharan (2006) argue for the use of laboratory experiments for investigating regulatory compliance theories given the paucity of reliable and available field data. Using an experimental design that allows identification of the effects of changing the audit probability, fixed audit cost, and marginal penalty, we broadly confirm the comparative statics of the theory. Further, the competitive endogenous audit mechanisms induce significantly higher compliance rates, and lower variation in compliance rates, than random audit regimes. In contrast, three previous experimental studies related to dynamic audits provide mixed evidence. Alm, Cronshaw, and McKee (1993) find that random audits result in higher compliance than do dynamic, "forward-looking" tax audits. Clark, Friesen, and Muller (2004) find that the mechanisms proposed by Harrington (1988) and Friesen (2003) lead to compliance rates that are no higher than random audits, although they increase audit efficiency. Cason and Gangadharan (2006) find moderate support for the theoretical predictions of Harrington's model, but do not include random audits as a comparative baseline.

## 2. Models

We frame our theory in the context of a regulation requiring disclosure of an activity which we will call the level of emissions. Later in this section, we discuss the applicability of these models to enforcement of regulations requiring actions such as pollution abatement. Disclosure of emissions is assumed to have a constant marginal cost, which could result from an emissions tax, but also could incorporate other costs such as those emanating from a negative market reaction. An audited firm pays a marginal penalty on emissions determined by the audit to have been unreported, and this penalty is assumed to be at least as high as the unit cost of disclosed emissions. The penalty represents any regulatory fines imposed, but again also may entail other costs to the firm of being found noncompliant with the disclosure requirement. Actual emissions  $e$  are exogenously determined; firms choose only how much to disclose. Audits are subject to error in quantifying a firm's emissions such that an audit reveals emissions of  $e + t$ , with  $t$  being drawn from the distribution  $F(t)$ , which is assumed to have positive density  $f(t)$  on the interval  $[a, b]$ . Note that we impose little structure on the distribution of audit errors at this point. If audit errors are one sided (meaning an audit cannot reveal emissions in excess of those actually emitted, so errors involve only failure to detect emissions), then  $a < 0$  and  $b = 0$ . If audits yield an unbiased estimate of emissions, then  $E[t] = 0$ . We will assume  $a > -e$  so that an audit cannot reveal negative emissions. We use the following notation:

$\alpha$  the cost to a firm of disclosed emissions ("tax")

$\beta$  the cost to a firm of revealed undisclosed emissions ("penalty")

$\gamma$  the cost to a firm of being audited

$e$  a firm's quantity of emissions

$q$  the quantity of emissions a firm chooses to disclose

We first describe firms' optimal disclosure choices under a random audit enforcement mechanism, and then develop two endogenous audit models. Both competitive endogenous mechanisms assume that the regulator compares firms' reported emissions and is more likely to audit a firm that fails to disclose more of its emissions. The first model, which we term a "tournament" mechanism because it is essentially a variant of the seminal Lazear and Rosen (1981) rank-order tournament game, assumes the regulator has a fixed audit capacity and therefore audits a certain number of firms, which is common knowledge. The second model is somewhat more general because it assumes simply that the firm's audit probability is a function of the difference between its degree of noncompliance and that of its peers. Although the probability each firm is audited depends on relative disclosure, whether each firm is audited is a result of an independent draw, which implies the regulator does not have fixed audit capacity. We term this the generalized relative evaluation (GRE) model. Models of this sort, where players' payoffs depend on the absolute difference between their effort and the average effort of competitors rather than rank, have received much less attention than rank-order tournaments, but Knoeber and Thurman (1994) discuss both types of models in the context of broiler (chicken) production.

**Random audit mechanism.** Suppose a firm is simply audited at random with probability  $p$  which is independent of whether other firms are audited. Employing a standard enforcement framework similar to that developed in Evans, Gilpatric, and Liu (2009), firm  $i$  chooses the optimal quantity of emissions to disclose to minimize its expected costs

$$(1a) \quad \min_{q_i} \alpha q_i + p \left\{ \gamma + \beta \int_{q_i - e_i}^b (e_i + t - q_i) f(t) dt \right\}.$$

So long as an interior solution exists, the optimal extent of over-/underreporting of emissions is independent of the actual quantity of emissions, and depends only on the regulatory parameters  $\alpha$ ,  $\beta$ , and  $p$  and on the distribution of audit errors  $F(t)$ . Define  $z_i \equiv q_i - e_i$ , so a negative  $z$  represents underreporting whereas a positive value represents overreporting. The reporting choice can then be restated as

$$(1b) \quad \min_{z_i} \alpha(e_i + z_i) + p \left\{ \gamma + \beta \int_{z_i}^b (t - z_i) f(t) dt \right\}.$$

The optimal choice of disclosure,  $z^*_i$ , is implicitly defined by

$$(2) \quad \frac{\alpha}{p\beta} = \int_{z_i^*}^b f(t) dt = 1 - F(z_i^*).$$

Under random audits, an interior solution exists for  $z^*_i$  on the interval  $[a, b]$  if  $0 < \alpha / p\beta < 1$ , with  $z^*_i$  defined by (2) above. For  $\alpha > p\beta$  it is not optimal to report any emissions, so a corner solution at  $q_i = 0$  obtains.<sup>3</sup> At an interior solution, the firm's optimal report is decreasing in the reporting cost, increasing in the probability of audit, and increasing in the penalty on revealed but unreported units (these results follow directly from the fact that  $F$  is an increasing function of  $z_i$ ). The solution is independent of the cost of being audited,  $\gamma$ . The characteristics of the enforcement regime determine the optimal extent of over- or underreporting, but this is independent of the emissions level.

**Tournament audit mechanism.** We now model  $N$  regulated firms in a peer group, which could represent an industry, geographical region, or other grouping. The regulator will audit  $k$  firms after receiving their disclosure reports. If audited, the situation is identical to that for the random audit mechanism, that is, firms face the cost in  $\{\cdot\}$  in expression (1). We assume the regulator audits those firms that it believes have most greatly underreported their emissions, but that making such judgments is prone to significant error, which introduces randomness into the selection process. Specifically, we assume the regulator has expectations regarding the level of emissions for each firm  $i$ , which we denote  $\bar{e}_i$ , representing the benchmark level of emissions against which a firm's report will be judged. Firms do not know the true value of the regulator's expectations, but believe them to be unbiased. From the firms' perspective let  $\bar{e}_i = e_i + \varepsilon_i$ , with  $\varepsilon_i$  being an *i.i.d.* draw for each firm from the distribution  $G$  representing the

regulator error in estimating firms' probable emissions. We assume  $G$  is symmetrically distributed around zero. The regulator will audit the  $k$  firms for which the difference between expected and reported emissions,  $(\bar{e}_i - q_i)$ , is greatest. As  $(\bar{e}_i - q_i) = (e_i + \varepsilon_i) - (z_i + e_i) = (\varepsilon_i - z_i)$ , this is equivalent to auditing the  $k$  firms for which  $(z_i - \varepsilon_i)$  is lowest. Define  $y_i = z_i - \varepsilon_i$  to be firm  $i$ 's error-adjusted compliance.

The randomness we assume to be present in the regulator's benchmarks reflects the reality that the firms in any peer group are likely to be highly heterogeneous, making relative evaluation problematic. For example, even if the firms operate similar production facilities, their actual emissions will of course differ, as will their emissions per unit of output. A regulator would nevertheless likely have some information on the relative scale of different firms and differences in production technologies and processes that would allow it to form expectations. We emphasize that the model does not require that a regulator can perfectly compare reports across myriad differentiated firms to determine which has exhibited the lowest level of compliance; in fact, the model relies on relative evaluation being problematic and prone to error. All that is necessary for the model to be representative in this regard is that comparison of reports across peer firms provides some information about which firms are likely to be most noncompliant. Put differently, the model simply assumes that, by choosing to audit those firms that stand out from their peers by disclosing low levels of emissions relative to expectations based on factors the regulator can observe, the regulator can achieve some correlation between the degree of noncompliance and the audit probability.

Within the peer group, firms are identical with respect to the distribution from which errors in the regulator's benchmark are drawn,  $G$ , and the distribution of errors in audits,  $F$ . Peer firms are also assumed to have identical regulatory cost parameters,  $\alpha$ ,  $\beta$ , and  $\gamma$ . This makes the firms symmetric competitors in the tournament game. Note that firms may differ in their actual emissions,  $e_i$ . Because of this, symmetric behavior in which all firms choose an identical level of over-/underreporting may imply different levels of disclosure,  $q_i$ .

We will identify the symmetric equilibrium of a compliance tournament game. Let the probability a particular firm  $i$  is audited, which depends on its own report and that of all other firms, be denoted  $p_i(z_i, z_{-i})$ , where  $z_{-i}$  denotes the common report of all other firms. For clarity, we first present this model for the case when the regulator only audits the firm with the lowest error-adjusted compliance, that is,  $k = 1$ , and then generalize it. For this case, we have

$$(3) \quad p_i(z_i, z_{-i}) = \int g(\varepsilon_i)(1 - G(\varepsilon_i + z_i - z_{-i}))^{N-1} d\varepsilon_i.$$

Of course, for any  $z_{-i}$ , the probability  $i$  is audited decreases with  $z_i$ :

$$(4) \quad \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} = -(N-1) \int g(\varepsilon_i)g(\varepsilon_i + z_i - z_{-i})(1 - G(\varepsilon_i + z_i - z_{-i}))^{N-2} d\varepsilon_i < 0.$$

As we are interested in the symmetric equilibrium of the game, the value of this expression when  $z_i = z_{-i}$  plays a critical role:

$$(5) \quad \left. \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}} = -(N-1) \int (g(\varepsilon_i))^2 (1 - G(\varepsilon_i))^{N-2} d\varepsilon_i.$$

We can now consider the optimization problem for a particular firm  $i$ :

$$(6) \quad \min_{z_i} \alpha(e_i + z_i) + p(z_i, z_{-i}) \left\{ \gamma + \beta \int_{z_i}^b (t - z_i) f(t) dt \right\}.$$

The first-order condition for minimizing expected cost is

$$(7) \quad \alpha + \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \left\{ \gamma + \beta \int_{z_i}^b (t - z_i) f(t) dt \right\} - p(z_i, z_{-i}) \beta \int_{z_i}^b f(t) dt = 0.$$

We obtain the unique symmetric Nash equilibrium of the game by evaluating (7) when  $z_i = z_{-i}$ . This gives us an expression that implicitly defines equilibrium disclosure in the tournament mechanism. We denote this expression  $V^T$ :

$$(8) \quad V^T = \alpha + \left( \left. \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}} \right) \left\{ \gamma + \beta \int_{z_i}^b (t - z_i) f(t) dt \right\} - \frac{1}{N} \beta \int_{z_i}^b f(t) dt = 0.$$

To generalize the model for  $k \geq 1$  audits, we simply need to identify  $p_i(z_i, z_{-i})$  and  $\partial p_i(z_i, z_{-i}) / \partial z_i$  for the general case. Here we follow Nalebuff and Stiglitz (1983), who identify the probability of a contestant ranking in each position within a general tournament. Let  $j$  denote player  $i$ 's rank up from the bottom (e.g.,  $j = 1$  denotes ranking last). Let  $p_{ij}(z_i, z_{-i})$  be the probability player  $i$  ranks in exactly the  $j$ th position. This probability is

$$p_{ij}(z_i, z_{-i}) = \int \frac{(N-1)!}{(N-j)!(j-1)!} g(\varepsilon_i) (G(\varepsilon_i + z_i - z_{-i}))^{j-1} (1 - G(\varepsilon_i + z_i - z_{-i}))^{N-j} d\varepsilon_i.$$

The probability that  $i$  is audited is then

$$p_i(z_i, z_{-i}) = \sum_{j=1}^k p_{ij}(z_i, z_{-i}).$$

From this we derive the effect of disclosure on the probability of a particular rank  $j$  at a symmetric point:

$$\left. \frac{\partial p_{ij}(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}} = \int \frac{(N-1)!}{(N-j)!(j-1)!} (g(\varepsilon_i))^2 \{ (1-G(\varepsilon_i))^{N-j-1} (g(\varepsilon_i))^{j-2} \} \\ \times \{ (j-1)(1-G(\varepsilon_i)) - (n-j)G(\varepsilon_i) \} d\varepsilon_i.$$

The symmetric equilibrium effect of disclosure on the probability that  $i$  is audited is then

$$\left. \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}} = \sum_{j=1}^k \left. \frac{\partial p_{ij}(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}}.$$

With this result, equation (8) as stated above defines equilibrium disclosure in the tournament mechanism with  $k$  audits.

*Proposition 1.* If an interior solution  $z^T$  representing a symmetric equilibrium exists under the tournament audit mechanism, this solution entails higher disclosure than under random audits for identical values of  $\alpha$  and  $\beta$ , and given equivalent audit probabilities,  $k/N = p$ .

*Proof.* See Appendix A.

Proposition 1 is a central result that the competitive incentive generated by the tournament structure yields greater disclosure *ceteris paribus*. The marginal benefit of disclosure is increased by the tournament mechanism because it consists of both the reduction of the expected penalty conditional on being audited (the last term in (8)) and also the reduction in the probability of being audited.

As is the case with random audits, disclosure increases with the penalty rate in the tournament mechanism, but this is now true both because a higher penalty makes losing the tournament and consequently being audited more costly and because the benefit of compliance arising from a lower expected fine conditional on being audited is larger. The comparative static with respect to the penalty is signed as follows:

$$\text{sign} \frac{\partial z^T}{\partial \beta} = - \frac{\partial V^T}{\partial \beta} = - \left\{ \left( \left. \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}} \right) \int_{z_i}^b (t - z_i) f(t) dt - \frac{1}{N} \int_{z_i}^b f(t) dt \right\} > 0.$$

Because the disclosure choice affects the probability of being audited in the tournament mechanism, the greater the cost of being audited,  $\gamma$ , the greater the equilibrium level of disclosure,  $\text{sign} \frac{\partial z^T}{\partial \gamma} = - \frac{\partial V^T}{\partial \gamma} = - \left( \left. \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}} \right) > 0$ .

Equilibrium disclosure also increases as the marginal effect of disclosure on the probability of being audited in equilibrium,  $\left. \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right|_{z_i=z_{-i}}$ , increases. For a given distribution, this effect grows as the variance of the errors declines. In other words, the greater the noise, the weaker the leverage achieved by the tournament mechanism. However, as in all tournament models, sufficient variance of the errors is required for the existence of the pure-strategy equilibrium. These aspects of the role of noise in tournaments have been examined by Lazear and Rosen (1981), Nalebuff and



Stiglitz (1983), and O'Keeffe et al. (1984). The effect of an increase in the number of competitors (or size of the peer group)  $N$ ,

$$\text{sign} \frac{\partial z^T}{\partial N} = -\frac{\partial V^T}{\partial N} = -\left\{ \frac{\partial \left( \frac{\partial p_i(z_i, z_{-i})}{\partial z_i} \right) \Big|_{z_i=z_{-i}}}{\partial N} \left\{ \gamma + \beta \int_{z_i}^b (t - z_i) f(t) dt \right\} + \frac{1}{N^2} \beta \int_{z_i}^b f(t) dt \right\},$$

is ambiguous because (holding the number of audits constant) this reduces the equilibrium audit probability for a firm (reducing the compliance incentive), but may increase the marginal effect of abatement on the probability of ranking last (i.e., it may increase the intensity of competition). The latter effect is complex, depending on the initial group size and on the error distribution (Nalebuff and Stiglitz, 1983, discuss the effect of group size on tournaments). Because the marginal effect of abatement on the probability of ranking last is unaffected by  $N$  when errors are uniform,  $\partial(\partial p_i(z_i, z_{-i}) / \partial z_i | z_i = z_{-i}) / \partial N = 0$ , this is a sufficient condition for an increase in  $N$  to reduce disclosure. However, from a regulator's perspective, it is important to recognize the possibility that increasing the group size (holding the number of audits constant) can actually increase disclosure in some circumstances if the error distribution has a strong central tendency which causes the increased competitive effect to dominate the effect of a lower equilibrium audit probability.

Compared to other rank-order tournament models ours is novel because although payoffs depend on rank, the expected payoff also varies conditional on rank as the expected penalty if audited falls with greater disclosure (compliance effort). Standard models assume payoffs are set in advance and depend solely on rank. Most tournament models are framed as competition to win (i.e., rank at the top), rather than to avoid ranking at the bottom, but this is not very critical to tournament theory. Its principal consequence, as Nalebuff and Stiglitz (1983) note, is that tournaments with penalties for ranking at the bottom rather than prizes for ranking at the top more easily satisfy the general incentive compatibility constraint, which is the condition that zero effort does not yield a higher payoff than playing the symmetric equilibrium when one's competitors play the symmetric equilibrium. One concern in tournament settings is the possibility of collusion, which can occur in equilibrium if a tournament is repeated indefinitely and players are sufficiently patient. By colluding, players in a tournament reduce their effort levels (disclosure in our model) and are all better off than at the symmetric equilibrium of the one-shot game. Note, however, that the optimal collusive behavior by regulated firms in this context would be to behave exactly as though the audit process were random, so even perfect collusion would not imply zero compliance (unless  $p\beta < \alpha$ ). Therefore, even if collusion undermined the leverage gained from this audit mechanism, it would never result in lower disclosure than with random enforcement. Although in an environment with heterogeneous players perfect collusion would be very difficult to achieve (requiring players to successfully coordinate), it remains an empirical question whether collusion would arise.

**Generalized relative evaluation audit mechanism.** A tournament is one particular type of competitive mechanism, or relative performance evaluation, in which a fixed number of competitors are ranked according to performance on a specified metric with a specified number of prizes or penalties determined by rank. This structure fits our regulatory enforcement setting well if the size of the peer group is known to all firms and the audit capacity of the regulator is fixed. However, relative performance evaluation may still be used by a regulator when the size of a peer group is not clearly defined and the audit capacity of the regulator is unknown to firms. In this section, we consider a less structured and therefore somewhat more general model in which the regulator compares firms' reports, and the probability a firm is audited is a function of how its report compares to the average of the other firms in the peer group.

We assume that for any firm  $i$  the probability of being audited is a function of the difference between  $i$ 's report and its actual emissions, that is,  $z_i$ , and the average of this difference among the other firms in the peer group. Define the average of  $i$ 's peers to be  $\bar{z}_i = \frac{1}{N-1} \sum_{j \neq i} z_j$ . The probability  $i$  is audited is then  $p(z_i, \bar{z}_i) = P(\bar{z}_i - z_i)$ . Note that in the tournament model, errors by the regulator in estimating firms' emissions and establishing a benchmark for comparison of compliance were explicitly modelled. Here, the audit probability function  $P(\bar{z}_i - z_i)$  implicitly captures the same notion that the regulator has sufficient information to make a firm's audit probability a decreasing function of its compliance relative to other firms, but errors in establishing benchmarks for comparison make this process imperfect so that it will not necessarily be the case that the least compliant firms are audited.

We assume  $\partial P / \partial z_i < 0$  for values which place  $P$  on the interval  $(0,1)$ . As with the tournament model, of particular importance is the point of symmetry where  $z_i = \bar{z}_i$ . We assume  $P(0)$  is strictly positive, so that in symmetric equilibrium each firm has a positive probability of being audited.

The cost minimization problem for a firm in this context is very similar to that in the tournament model, with the form that  $p(z_i, \bar{z}_i)$  takes being the only difference, so we simply state the condition for equilibrium disclosure:

$$(9) \quad V^{GRE} = \alpha + \frac{\partial P(0)}{\partial z_i} \left\{ \gamma + \beta \int_{z_i}^b (t - z_i) f(t) dt \right\} - P(0) \beta \int_{z_i}^b f(t) dt = 0.$$

*Proposition 2.* If an interior solution  $z^{GRE}$  representing a symmetric equilibrium exists under the GRE audit mechanism, this solution entails higher disclosure than under random audits for identical values of  $\alpha$  and  $\beta$ , and given equivalent equilibrium audit probabilities,  $p = P(0)$ . The GRE and tournament mechanisms yield identical disclosure for identical values of  $\alpha$  and  $\beta$ , given equivalent equilibrium audit probabilities,  $k/N = P(0)$ , and given equal marginal effects of disclosure on the probability of being audited,  $\partial p_i(z_i, \bar{z}_i) / \partial z_i |_{z_i = \bar{z}_i} = \partial P(0) / \partial z_i$ .

*Proof.* See Appendix A.

Proposition 2 states that the GRE mechanism, like a tournament, achieves leverage relative to random audits by inducing competition to avoid being audited. Furthermore, if in equilibrium the marginal effect of a firm's compliance effort on the audit probability and the audit probability itself are identical between a tournament and the GRE, then the two mechanisms yield equivalent equilibrium compliance (disclosure). It should be noted, however, that away from the symmetric equilibrium the two mechanisms may differ. In a tournament the marginal effect on the probability of being audited,  $\partial p_i(z_i, z_{-i}) / \partial z_i$ , is not linear in its arguments and peaks at the symmetric point. Therefore, when firms are not behaving symmetrically, the competitive incentive is weakened. Although an equivalent, nonlinear audit function can be used as part of the GRE, marginal incentives can be manipulated. For example, if  $P(z_i - z_{-i})$  is a linear function, then the mechanism maintains a constant marginal effect of compliance effort on a firm's audit probability so long as a corner is not reached (with the audit probability 0 or 1), so the competitive incentive is not weakened by asymmetric behavior among firms. Also, as discussed above, an important difference between these two competitive endogenous audit mechanisms is that the GRE mechanism is not sensitive to group size,  $N$ . Group size has no effect on the predicted equilibrium because it does not change either the equilibrium audit probability or the marginal effect of compliance effort.

**Applicability to enforcement of regulated actions.** The competitive endogenous audit mechanisms we have developed are framed in an information disclosure framework; the regulated firm must disclose its performance, such as emissions released. Regulations stipulating actions such as pollution abatement are less clearly amenable to this type of enforcement, because they do not involve firms making a public signal (the report) to the regulator which can be used to condition audits. Nevertheless, the framework here can apply if the regulator can costlessly obtain some signal (however noisy) of firms' relative compliance effort prior to choosing which firms to audit. For example, for some types of pollution regulations this could involve observing changes in ambient air or water quality in the vicinity of firms' plants, or could simply involve being in contact with industry insiders who may have a sense of which firms are particularly focused on regulatory compliance and which are not. As we have discussed, competitive incentives arise whenever firms' probability of being audited becomes conditional on their relative compliance effort, and this will occur if a regulator can in any way make the selection of firms for audit correlated with relative compliance effort.

### 3. Experimental design

The main objectives of the compliance experiment are to test the main comparative statics of the theory and examine the leverage gained through competitive endogenous audits. The experiment instructions use neutral framing (instructions included in Appendix B), but for ease of exposition we describe the experimental design and discuss results using the environmental disclosure context. The compliance experiment involves sessions of 20 players who participate in groups of  $N = 5$  over two sequences of 20 decision periods. Players are randomly rematched into groups prior to each period to preserve as well as possible the incentives of the one-shot game while allowing for

learning. The actual number of periods is undisclosed in order to minimize possible end-of-game effects. In each period, players receive endowment  $E$  and produce a fixed amount of emissions (“actual output” in the experiment),  $e$ , of 20.6 The decision task for each player is to choose a level of disclosure (“reported output”), at a per-unit tax (“reporting cost”) of 1 experimental dollar, by selecting an integer between 0 and 40 (inclusive). After all choices are made, one or more players are selected for audit (“inspection”). Players are made aware that, if audited, the process is imperfect in that it can reveal an emissions level that differs from their actual emissions of 20.

Each player faces one of the three audit mechanisms discussed in the theory section: random audit, tournament audit, and GRE audit. Under random audit, players are randomly selected according to a specified probability (“chance”). For the tournament audit, the reports are error

**TABLE 1 Experiment Parameters**

Parameter/Variable	Description	Value(s)
$\beta$	Penalty	1 or 3
$\gamma$	Audit cost	0 or $\frac{40}{3}$
$p$	Audit probability	20% or 60%
$\alpha$	Tax	1
$e$	Emissions	20
$t$	Audit error	<i>i.i.d.</i> Uniform[−20, 20]
$s$	Regulator error	<i>i.i.d.</i> Uniform[−10, 10]

adjusted (the “computer” makes an “initial estimate” of “actual output”) by adding to each disclosure report a mean-zero, *i.i.d.* draw from the uniform distribution with supports [−10, 10]. Then, the  $k$  player(s) in the group with the lowest error-adjusted report(s) is selected for audit. Note that with the distribution from which errors are drawn being uniform, the effect of disclosure on the audit probability is greatly simplified compared to the general case of  $k$  audits developed in Section 2. With uniformly distributed errors, only the probabilities of ranking first and last are affected by effort at a symmetric point; the probability of all intermediate ranks is unchanged. This in turn means the effect of effort on the probability of being audited is independent of  $k$  and equal to the effect of effort on the probability of ranking first, as stated in equation (5).

Under the GRE mechanism, a player’s level of disclosure relative to the average of other group members determines the audit probability. In particular, the audit probability equals

$$P_i = \begin{cases} P(0)(\bar{z}_{-i} - z_i) & \text{if } (\bar{z}_{-i} - z_i) \in [-P(0), 1 - P(0)] \\ 1 & \text{if } (\bar{z}_{-i} - z_i) > 1 - P(0) \\ 0 & \text{if } (\bar{z}_{-i} - z_i) < -P(0). \end{cases} \quad (11)$$

This yields a symmetric equilibrium audit probability of  $P(0)$ , and a marginal effect of disclosure on the probability of being audited in equilibrium of −1, which is consistent

with the tournament mechanism. Information on the GRE mechanism is provided through an (simpler) equation, a “Chance of Inspection Table” (see Appendix B), which displays the audit probability associated with various possible disclosure outcomes, and descriptive text.

For players selected for audit, they (possibly) pay a fixed audit cost (“inspection cost”). The audit is unbiased and reveals an estimate of emissions (“estimated output”) by adding to actual emissions a random variable,  $\varepsilon_i$ , which is *i.i.d.* uniform with supports  $[-20, 20]$ . A penalty is levied on any (estimated) undisclosed emissions. The feedback given at the end of the decision period includes: (i) whether the player was audited, and if so estimated emissions; (ii) all relevant earnings calculations; and (iii) the disclosure reports of all group members and whether they were audited.

Each of the two 20-period sequences involves a different parameter combination or “treatment.” Overall, there are eight treatments, which represents a full factorial of the following attributes: penalty ( $\beta = 1$  and  $\beta = 3$ ); audit cost ( $\gamma = 0$  and  $\gamma = 40/3$ ); and (expected) audit probability ( $p = 0.2$  and  $p = 0.6$ ). Table 1 summarizes the design parameters of the experiment. The lone parameter that varies within session is the audit cost, and the order in which the two audit costs are encountered varied across sessions so that order effects can be sufficiently controlled for through statistical modelling.<sup>7,8</sup> For the tournament audit, given five-player groups,  $k = 1$  and  $k = 3$  are chosen in order to correspond to audit probabilities of 0.2 and 0.6. In the GRE sessions, the audit probabilities are attained by setting  $p = P(0)$ .

**Testable hypotheses.** The experimental design lends itself to testing the theory in several ways. First, six of the eight treatments represent cases where there is theoretically zero disclosure under random audits (given  $\alpha > p\beta$ ), but nevertheless positive disclosure with the competitive endogenous audits. This thus provides a basis to explore the leverage afforded by competitive endogenous audits. Second, the GRE is parameterized to generate identical predictions to the tournament audit for each treatment. Third, common to related tax and environmental compliance experiments, the chosen audit probabilities are much larger than those in relevant, naturally occurring situations for purposes of transparency and saliency. For the competitive endogenous audit mechanisms, the predicted marginal effect of changing the probability from 20% to 60% is an increase in the disclosure *rate* between 0.2 and 0.3.<sup>9</sup> Fourth, in a similar vein, predicted disclosure rates for the competitive endogenous audit mechanisms range from 0.2 to 1.7. There are meaningful differences between any treatment pair, and predicted undercompliance, approximate compliance, and overcompliance. Fifth, the design allows us to test for equivalence between two components of the audit the regulator has control over: the penalty and audit cost. In particular, under competitive endogenous audits, the combination  $\{\beta = 1, \gamma = 40/3\}$  yields the same disclosure prediction as the combination  $\{\beta = 3, \gamma = 0\}$  for a given audit probability. The main hypotheses to be tested are summarized below

Hypothesis 1. Competitive endogenous audits lead to higher disclosure than random audits.

Hypothesis 2. Tournament and GRE audits generate equivalent disclosure.  
Hypothesis 3. Random audits: disclosure is invariant to audit cost.  
Hypothesis 4. Competitive endogenous audits: increasing the penalty, audit cost, or audit probability increases disclosure.  
Hypothesis 5. Competitive endogenous audits: increasing the penalty or audit cost has an equivalent effect on disclosure.

**Participant pool and procedures.** Two hundred and forty undergraduate students at the University of Tennessee participated in the experiment during the spring of 2009. We conducted 12 sessions, and each session included 20 participants who played in two treatments of the same audit mechanism. Participants were drawn from a large pool of students who had previously registered to be potential participants in economics experiments. The participant pool is similar to the general undergraduate population of the university in terms of age, gender, and academic major. Experiments were conducted in a designated experimental laboratory. Participant earnings were denominated in experiment currency, which was exchanged for dollars at the end of the session at the known rate of 20 to \$1 (U.S.). The experiment lasted approximately 90 minutes and individual earnings averaged \$26.

Decisions were made via networked computers using software programmed with z-Tree (Fischbacher, 2007). The software collected all decisions and made earnings calculations. Written instructions were provided to each participant, which were read aloud by the same author in each session. To help facilitate learning, participants were then asked to work through three calculation questions (using pencil and paper) and were paid \$1 for providing correct answers. The questions involved arbitrarily making a disclosure choice and then determining earnings under three audit outcomes. Experiment moderators privately checked the calculations and explained procedures in the case of wrong answers. Prior to paid decision periods, there were three practice periods. At the conclusion of the experiment, a short questionnaire was administered that included among other things an assessment of how well instructions were understood as well as demographic questions.

#### 4. Results

**Analysis of means.** To facilitate tests of the hypotheses outlined above, we estimate a linear regression model of disclosure rates that allows identification of mechanism-specific treatment main effects.<sup>10,11</sup> As stated above, the only treatment variable that varied within subject is  $\gamma$  (audit cost). Simple examination of the raw data suggests prominent order effects as it pertains to this parameter. Further, there is evidence of learning. This is most evident in random-audit sessions with predicted disclosure rates of zero, as subjects tended to overreport relative to theoretical predictions and then decreased disclosure after repetition. Based on tests for structural breaks, this trending behavior flattens by period 15 for all treatments.

Motivated by these observations, we control for learning and order effects in the econometric model. As a transparent control for learning, we estimate separate main effects for two period groupings, periods 1–15 and periods 16–20. We created six

variables to control for order effects. These variables are interactions between an indicator variable, which equals 1 if a treatment was encountered in the second sequence in a session, with each treatment effect (*High  $\beta$* , *High  $\gamma$* , and *High  $p$* ) and with each mechanism.

Formally, the treatment main-effects model is given by

(11)

$$r_{ij} = \sum_{m=1}^3 \sum_{n=1}^2 I_{ij}^{m,n} \cdot (\theta_0^{m,n} + \theta_1^{m,n} \cdot (\text{High } \beta)_{ij} + \theta_2^{m,n} \cdot (\text{High } \gamma)_{ij} + \theta_3^{m,n} \cdot (\text{High } p)_{ij}) + \varphi C_{ij} + u_{ij},$$

where  $r_{ij}$  is the disclosure rate, that is, the share of emissions disclosed, for individual  $i$  in period  $j$ ;  $I_{ij}^{m,n}$  is an indicator variable that takes a value of 1 if the disclosure rate corresponds to mechanism  $m$  and period grouping  $n$  and equals 0 otherwise, where  $n = 1$  denotes periods 1–15 and  $n = 2$  denotes periods 16–20; *High  $\beta$* , *High  $\gamma$* , and *High  $p$*  are indicator variables that equal 1 to denote treatment parameters  $\beta = 3$ ,  $\gamma = 40/3$ , and  $p = 0.6$ , respectively, and equal 0 otherwise;  $C_{ij}$  denotes the vector of control variables related to order effects;  $\theta$  and  $\varphi$  are parameter vectors to be estimated; and  $u_{ij}$  is a mean zero error term.

To freely allow model errors to be autocorrelated and conditionally heteroskedastic for an individual, as well as to allow this error correlation to vary across all individuals, we use robust standard errors adjusted for clustering at the individual level (i.e., “cluster-robust” standard errors) and associated robust test statistics. Errors across individuals and groups are assumed to be independent. This covariance estimator is a generalization of the oft-used heteroskedasticity consistent covariance estimator of White (1980), and is similar to the Newey and West (1987) HAC covariance estimator except that all autocovariances particular to an individual are included and no kernel weighting function is used. Monte Carlo evidence suggests that test statistics based on this covariance estimator have the correct size for panel data with a moderate number of cross-sectional units under various data-generating processes (Bertrand, Duflo, and Mullainathan, 2004; Kezdi, 2004; Vossler, 2008).

Table 2 presents selected estimation results, in particular the treatment main effects corresponding to the period grouping 16–20. Based on the estimated coefficients, we generated treatment-specific means for each mechanism and present these along with the Nash equilibria in Table 3 and Figure 1. Furthermore, Table 4 includes tests of treatment differences across mechanisms. These tests lead us to two of our main empirical results.



**TABLE 2 Treatment Main-Effects Model**

Dependent Variable: Disclosure Rate

Variable	Description	Coefficient Estimates		
		Random	Tournament	GRE
<i>High <math>\beta</math></i>	= 1 if $\beta = 3$ ; = 0 if $\beta = 1$	0.298* (0.060)	0.490* (0.073)	0.311* (0.070)
<i>High <math>\gamma</math></i>	= 1 if $\gamma = \frac{40}{3}$ ; = 0 if $\gamma = 0$	0.056 (0.062)	0.453* (0.065)	0.361* (0.066)
<i>High <math>p</math></i>	= 1 if $p = 0.6$ ; = 0 if $p = 0.2$	0.399* (0.061)	0.376* (0.051)	0.288* (0.044)
Constant		0.004 (0.060)	0.208* (0.069)	0.403* (0.043)
Controls for order effects?			Yes	
Controls for learning?			Yes	
$R^2$			0.88	
$F$			331.73*	
$N$			9600	

**TABLE 3 Nash Equilibrium and Mean Disclosure Rate by Mechanism and Treatment**

	Nash Equilibrium		Mean Disclosure Rate		
	Random Audit	Tournament/GRE	Random Audit	Tournament	GRE
Treatment 1: $\beta = 1, \gamma = 0, p = 0.2$	0.00	0.19	0.004 (0.060)	0.208 (0.069)	0.403* (0.039)
Treatment 2: $\beta = 1, \gamma = 0, p = 0.6$	0.00	0.51	0.402* (0.079)	0.583 (0.076)	0.691* (0.050)
Treatment 3: $\beta = 1, \gamma = \frac{40}{3}, p = 0.2$	0.00	0.51	0.060 (0.033)	0.661* (0.046)	0.764* (0.068)
Treatment 4: $\beta = 3, \gamma = 0, p = 0.2$	0.00	1.03	0.302* (0.056)	0.698* (0.036)	0.714* (0.067)
Treatment 5: $\beta = 1, \gamma = \frac{40}{3}, p = 0.6$	0.00	1.03	0.458* (0.059)	1.037 (0.067)	1.052 (0.068)
Treatment 6: $\beta = 3, \gamma = 0, p = 0.6$	0.89	1.30	0.701* (0.065)	1.073* (0.041)	1.002* (0.077)
Treatment 7: $\beta = 3, \gamma = \frac{40}{3}, p = 0.2$	0.00	1.50	0.358* (0.060)	1.151* (0.066)	1.075* (0.041)
Treatment 8: $\beta = 3, \gamma = \frac{40}{3}, p = 0.6$	0.89	1.70	0.757 (0.066)	1.527* (0.077)	1.363* (0.049)

Result 1. All comparative statics of the three audit mechanisms are confirmed.

Result 2. (i) Disclosure rates with competitive endogenous audits are higher than with random audits. (ii) Furthermore, with little qualification, tournament and GRE audits lead to similar disclosure rates.

Result 1 is directly evident from the estimated treatment main effects presented in Table 2. Consistent with Hypothesis 3, in the random audit the marginal effect of increasing the audit cost from  $\gamma = 0$  to  $\gamma = \frac{40}{3}$  on the disclosure rate is just 0.056, which is not statistically



FIGURE 1 DISCLOSURE RATES BY MECHANISM AND TREATMENT

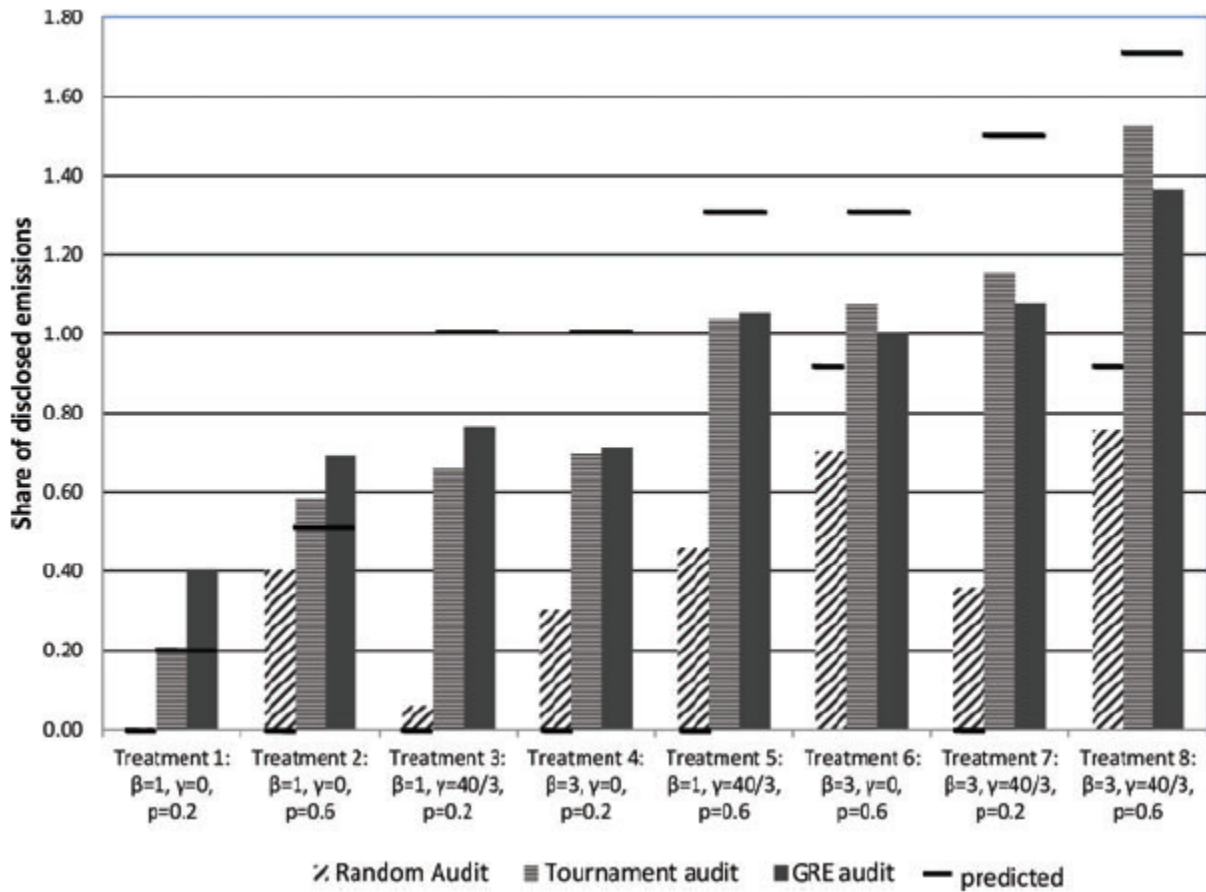


TABLE 4 Difference of Means Tests; Difference of Variances Tests

	H <sub>0</sub> : Equal Disclosure Rate			H <sub>0</sub> : Equal Variance		
	Tourn. = Random	GRE = Random	Tourn. = GRE	Tourn. = Random	GRE = Random	Tourn. = GRE
Treatment 1: $\beta = 1, \gamma = 0, p = 0.2$	0.204* (0.051)	0.400* (0.074)	-0.196* (0.082)	0.033 (0.024)	-0.032 (0.026)	0.065* (0.030)
Treatment 2: $\beta = 1, \gamma = 0, p = 0.6$	0.181* (0.079)	0.289* (0.093)	-0.108 (0.089)	0.017 (0.034)	-0.094* (0.028)	0.111* (0.034)
Treatment 3: $\beta = 1, \gamma = \frac{40}{3}, p = 0.2$	0.601* (0.057)	0.704* (0.071)	-0.103 (0.081)	0.081* (0.030)	-0.003 (0.031)	0.084* (0.036)
Treatment 4: $\beta = 3, \gamma = 0, p = 0.2$	0.396* (0.066)	0.412* (0.069)	-0.016 (0.076)	-0.131* (0.028)	-0.065* (0.030)	-0.066* (0.032)
Treatment 5: $\beta = 1, \gamma = \frac{40}{3}, p = 0.6$	0.578* (0.088)	0.594* (0.083)	-0.015 (0.094)	0.065 (0.049)	-0.065* (0.034)	0.130* (0.051)
Treatment 6: $\beta = 3, \gamma = 0, p = 0.6$	0.373* (0.076)	0.302* (0.087)	0.071 (0.085)	-0.147* (0.033)	-0.127* (0.042)	-0.020 (0.038)
Treatment 7: $\beta = 3, \gamma = \frac{40}{3}, p = 0.2$	0.793* (0.067)	0.717* (0.071)	0.076 (0.076)	-0.083* (0.025)	-0.036 (0.027)	-0.047 (0.028)
Treatment 8: $\beta = 3, \gamma = \frac{40}{3}, p = 0.6$	0.770* (0.083)	0.606* (0.082)	0.164 (0.091)	-0.099* (0.043)	-0.098* (0.041)	-0.001 (0.047)

different from zero.<sup>14</sup> For the competitive endogenous audit mechanisms, all treatment effects are positive and statistically significant, which confirms Hypothesis 4. Further, the marginal effects of the penalty and audit cost are similar in size and are statistically equivalent, consonant with Hypothesis 5: tournament audit (0.490 versus 0.453;  $p$  value = 0.55); GRE audit (0.311 versus 0.361;  $p$  value = 0.68).

Result 2 is best illustrated with Figure 1. In particular, the mean disclosure rates are systematically higher with the competitive endogenous mechanisms under all treatment conditions. At the same time, the tournament and GRE audits have similar disclosure rates across treatment conditions. The ocular patterns are verified with the statistical tests presented in Table 4. In all cases, estimated mean disclosure rates are statistically higher for the competitive endogenous mechanisms relative to random audits. Further, although there is a tendency for tournament audits to motivate lower disclosure levels under low and moderate treatment conditions, and higher relative disclosure under more draconian parameters, under all treatment conditions the two competitive endogenous audit mechanisms lead to statistically identical disclosure rates with the exception of  $\{\beta = 1, \gamma = 0, p = 0.2\}$ .

Result 3. For all audit mechanisms, when predicted disclosure rates are less than 1, there is a tendency of overcompliance; otherwise, when predicted disclosure rates are equal to or greater than 1, there is undercompliance.

In experimental tax compliance studies, under design conditions that lead to theoretical predictions of less than 100% compliance, compliance rates above those predicted by theory have been widely observed (e.g., Alm, McClelland, and Schulze, 1992). To the best of our knowledge, no related study on compliance has examined conditions under which theory predicts perfect compliance or overcompliance. As seen in Figure 1 and supported by statistical tests, in all such cases estimated disclosure rates are statistically different from predicted and are substantially less. Thus, although too much disclosure—under theoretical conditions of undercompliance— may be supported by social norms or the overweighting of small probabilities (Alm, McClelland, and Schulze, 1992), these or other factors seemingly lead to less than predicted disclosure under conditions that should lead to exact or overreporting.

**Analysis of variances.** In laboratory experiments involving tournaments, a common finding is that there is a large variance in individual-level decisions relative to other incentive mechanisms such as piece rates and collective compensation schemes (e.g., Bull, Schotter, and Weigelt, 1987; Nalbantian and Schotter, 1997; van Dijk, Sonnemans, and van Winden, 2001). As discussed in this literature, this may be due to the presence of strategic incentives in tournaments as well as increased computational difficulty. There are similar considerations when comparing random audits with the competitive endogenous mechanisms. To explore this issue, we estimate the model defined by (11) but using instead as the dependent variable the squared deviation from the mean disclosure rate, that is,  $(r_{ij} - \bar{r}_j)^2$ , where  $\bar{r}_j$  is the treatment-specific mean in period  $j$ . Thus, a measure of zero suggests that all players in a particular treatment made the same decision.

The last three columns in Table 4 present pairwise tests of equal variances. In just 1 of the 16 comparisons (Treatment 3, tournament), we find that the variance is significantly different and lower with random audits. In fact, in 9 cases, random-audit mechanisms are associated with a *higher* variance in disclosure rates.<sup>15</sup> To give these results some perspective, across all treatments, the random audit variance is 2.2 times that of tournament audits and 1.4 times that of the GRE. Thus, the general finding is that, in contrast to experimental findings from related studies, the competitive incentives introduced by the endogenous audit mechanisms actually lead to relatively lower variance in choices. Although the underlying mechanism for the disparity is not clear, it may be due to a difference in framing, that is, in the regulation setting, firms are competing not to lose rather than competing to win.

These findings are especially strong given that the comparisons are biased toward finding lower variance with random audits. This is because random audits generate lower disclosure—recall that predicted disclosure is zero in six of eight treatments—and players are hence constrained from disclosing amounts that are significantly less than average, as negative disclosure is not possible.

We further find that in half of the cases, the variance from the tournament mechanism is higher than the GRE, and this result goes in the opposite direction for Treatment 4. This weak evidence is consistent with the marginal compliance incentives when players act asymmetrically. In particular, as noted in the theory section, the GRE mechanism maintains a constant marginal effect of compliance effort on a firm's audit probability, whereas the tournament has nonlinear marginal incentives that are their highest at the symmetric Nash equilibrium.

**Robustness checks.** To evaluate the robustness of our results, we undertook additional analyses using the disclosure and disclosure variance data. In the first set of checks, using the same data set, we estimate three alternative models: one that only controls for order effects; one that only controls for learning effects; and a third that controls for neither, that is, a regression on treatment-specific indicator variables. For analyses of disclosure rates, with very slight qualifications, our main results hold based on these models. The qualifications are that Results 2(i) and 2(ii) only continue to hold on average (i.e., if we pool disclosure rates across treatments for a mechanism), but do not always hold when comparing across mechanisms on a treatment-by-treatment basis.

In the case of Result 2(i), this no longer holds for only Treatment 2, where for the random-audit mechanism we observed the sharpest time trend (toward zero). Turning to the analysis of the variance of disclosure rates, the comparison between competitive endogenous audits and random audits is virtually unaffected. However, the systematic differences between tournament and GRE variances become much less pronounced, and in the vast majority of cases we fail to reject the null hypothesis of equal variances.

As an additional robustness check, we estimated the models while restricting the data to include only those treatments that appeared as a first sequence in a session. This avoids completely the issue of order effects. Further, controls for learning effects were excluded. With this subset of the data, we continue to confirm the comparative statics results and find that the competitive endogenous audit mechanisms yield higher disclosure rates. Further, in 9 out of 16 cases, we find a lower variance in competitive endogenous audit mechanisms relative to random audits (and null effects in the remaining cases).

## **5. Conclusion**

The models developed here show that if a regulator evaluates firms' relative disclosure when selecting firms for audit, in an environment where both the selection process and the audits themselves are subject to errors, this naturally generates competition in compliance effort among regulated firms. Compliance effort reduces both a firm's expected penalty, conditional on being audited, and the probability that the firm will be audited. As our experimental results confirm, the addition of this second incentive can be quite powerful, especially when the cost of being audited is significant.

It is informative to compare the competitive endogenous audit mechanisms described here with the targeting framework developed by Harrington (1988) and others which conditions a firm's audit probability on past compliance rather than as a signal of relative current compliance effort. The targeting framework does not depend on a regulator having a contemporaneous signal of firms' relative compliance effort as do our models, and such a signal may be unobtainable in enforcement of mandated actions. However, targeting has some significant disadvantages. In particular, by its nature, similar firms (typically modelled as identical) are at any moment in time subject to differing audit probabilities, and thus choose different compliance efforts. In the dichotomous framework, this essentially means that leverage is achieved when only those firms that are targeted in any given period actually comply, whereas nontargeted firms do not comply. This does a poor job of explaining the stylized fact that most firms are compliant most of the time, and is not very satisfactory as prescriptive policy as it implies that targeting is effective only in circumstances when noncompliance by most firms at any given time is acceptable. Similarly, if compliance effort is a continuous rather than a dichotomous choice, targeting results in much higher compliance effort among targeted firms. This is inherently inefficient, given standard assumptions about the convexity of the cost of compliance effort that imply that the costs of achieving any given level of aggregate industry compliance are minimized by equalizing compliance effort among firms.

An important advantage of the competitive endogenous audit mechanisms we have developed is that they achieve leverage relative to random audits without dividing firms into groups with different enforcement intensity and compliance incentives. Both targeting and competitive endogenous auditing use the cost of being audited to leverage enforcement by making the probability this cost is incurred dependent on firms' behavior. However, the competitive endogenous audit mechanisms apply enforcement "pressure" uniformly across all firms.

## Appendix A

### □ Proofs of Propositions 1 and 2

*Proof of Proposition 1.* We defined  $V^T$  as the expression that must equal 0 to satisfy the first-order condition for equilibrium in the tournament mechanism. Define the analogous expression for the random-audit mechanism  $V^{RA}$ :

$$V^{RA} = \alpha - p\beta \int_{z_t}^d f(t) dt.$$

Subtracting  $V^{RA}$  from  $V^T$  given identical values of  $\alpha$  and  $\beta$ , and  $p = k/N$ , yields the following expression:

$$V^T - V^{RA} = \left( \frac{\partial p_t(z_t, z_{-t})}{\partial z_t} \Big|_{z_t = z_{-t}} \right) \left\{ \gamma + \beta \int_{z_t}^d (t - z_t) f(t) dt \right\} < 0.$$

The fact that this difference is negative shows that the  $z_t$  that solves the cost minimization problem under random audits is lower than the equilibrium  $z_t$  under tournament audits.

*Proofs of Proposition 2.* The result that the GRE mechanism yields higher disclosure than the random-audit mechanism is analogous to Proposition 1. Subtracting  $V^{RA}$  from  $V^{GRE}$  given identical values of  $\alpha$  and  $\beta$ , and  $p = P(0)$ , yields the following expression:

$$V^{GRE} - V^{RA} = \left( \frac{\partial P(0)}{\partial z_t} \right) \left\{ \gamma + \beta \int_{z_t}^d (t - z_t) f(t) dt \right\} < 0,$$

which establishes that the  $z_t$  that solves the cost minimization problem under random audits is lower than the equilibrium  $z$  under GRE audits. Given identical values of  $\alpha$  and  $\beta$ , and  $k/N = P(0)$ , then

$$V^{GRE} - V^T = 0.$$

This establishes that the  $z_t$  that solves the cost minimization problem under the tournament mechanism is the same  $z_t$  that solves the cost minimization problem under GRE audits given the parameter restrictions stated.

## Appendix B

### □ Experiment instructions (GRE audit with $\beta = 3$ , $\gamma = 40/3$ , $p = 0.2$ )

#### INTRODUCTION

This experiment is a study of group and individual decision making. The amount of money you earn depends on the decisions that you make and thus you should read the instructions carefully. The money you earn will be paid privately to you, in cash, at the end of the experiment. A research foundation has provided the funds for this study.

In this experiment, you will be in a group consisting of five players: you and four others. The other players in your group are people sitting in this room, but you will not be told who is in your group.

You will make decisions privately, that is, without consulting other group members. Please do not attempt to communicate with other participants in the room during the experiment. If you have a question as we read through the instructions or at any time during the experiment, please raise your hand and an experiment moderator will come by to answer it.

The experiment is broken up into many decision “periods”. With the exception of your decisions in three practice periods, you will be paid based on your decision in each and every period. In other words, each decision you make is important in determining the amount of money you earn.

Your group will change from one period to the next. In particular, prior to each period, the computer randomly places players into groups.

Your earnings in the experiment are denominated in experimental dollars, which will be exchanged at a rate of 20 to \$1 U.S. at the end of the experiment.

## Overview

You are given initial earnings of \$36 each period.

In each decision period, consider yourself as having produced 20 units of output. Your decision is how much output to report.

You pay a cost based on how much output you report.

How much output you report affects the chance that your report will be inspected. If your report is inspected, you face additional costs.

Each decision period is independent from the others, in the sense that your earnings in one period depend only on your decision – and the decisions of others – in that particular period only.

## Your reporting decision

Your **actual output** each and every period is 20 units. Your sole decision is to choose what output to report. Your **reported output** can be any amount between, and including, 0 and 40.

For *each* unit of reported output, you pay a cost of \$1. We refer to this amount as your **reporting cost**.

## Determining who is inspected

Your chance of being inspected is based on the difference between your reported output and the average reported output of the other four players in your group. In particular, the computer determines your chance of inspection using the following relationship:

$$\text{Inspection chance} = 20\% + 5\% \times [\text{average report from the other players} - \text{your reported output}]$$

This information is further presented in the “**Chance of Inspection Table**”. As you can see from this table: (1) you face a **20%** chance of inspection if your reported output is the same as the average reported output of your other group members; (2) your chance of being inspected is greater than 20%, and increases, the further your reported output is below the average reported output of the other players in your group; and (3) your chance of being inspected is less than 20%, and decreases, the further your reported output is above the average reported output of the other players in your group.

## What happens if you get inspected

If you are inspected you pay an **inspection cost** of \$13.33. This cost is the same regardless of your reported output.

In addition, you may pay a **penalty**, which does depend on your report.

The computer makes an estimate of your output. This estimate is based on your actual output.

In particular, **estimated output** is equal to your **actual output** plus a random amount. The random amount has an equal chance of being any number between, and including, –20 and 20. The random amount is equal to 0 on average, which means on average estimated output is equal to your actual output. Another way to view the estimated output is that it has an equal chance of being any number between 0 and 40.

If the estimated output is *greater than* your reported output, you pay \$3 for each unit of output you are estimated to have under-reported. Otherwise, you do not pay any penalty.

## Your earnings

As described above, you are given initial earnings of \$36, and your overall earnings for the decision period depend upon how much you report (reporting cost) and – if you are inspected – an inspection cost and possibly a penalty.

Thus, after you have submitted your report, three things can happen: (1) You are not inspected; (2) You are inspected and your estimated output is *less than* your reported output; or (3) You are inspected and your estimated output is *greater than* your reported output. We summarize below how your earnings will be calculated under each scenario.



**Your earnings (You are not inspected)**

Since you are not inspected, you do not pay an inspection cost or a penalty.

Your earnings for the period are your initial earnings minus your reporting cost.

In particular:

	\$36	(Initial earnings)
–	Reported output × \$1	(Reporting cost)
–	\$0	(Inspection cost)
–	0	(Penalty)
<hr/>		
=	Period Earnings	

**Your earnings (You are inspected and your estimated output is *less than* your reported output)**

Since you are inspected you pay the inspection cost.

Since your estimated output is *less* than your reported output you do **not** pay a penalty.

Your earnings for the period are your initial earnings minus your reporting cost.

In particular:

	\$36	(Initial earnings)
–	Reported output × \$1	(Reporting cost)
–	\$13.33	(Inspection cost)
–	0	(Penalty)
<hr/>		
=	Period Earnings	

**Your earnings (You are inspected and your estimated output is *greater than* your reported output)**

Since your estimated output is greater than your reported output you pay a penalty of \$3 for each unit you are estimated to have under-reported. Your earnings for the period are your initial earnings minus your reporting cost, inspection cost and penalty.

In particular:

	\$36	(Initial earnings)
–	Reported output × \$1	(Reporting cost)
–	\$13.33	(Inspection cost)
–	[Estimated output – reported output] × \$3	(penalty)
<hr/>		
=	Period Earnings	

**Results**

After all the members of your group have made their decisions, you will see a results screen. This screen displays your reported output, whether you were inspected, and your earnings.

You will also see the reported output of all other players in your group and whether they were inspected.

**Example.** In this example you will calculate what your earnings for a decision period would be based on a particular scenario. If all of your calculations are correct we will give you US\$1.

Please choose your reported output: \_\_\_\_\_

Scenario A. Suppose you do **not** get inspected. Please calculate what your earnings would be based on your choice of reported output and write this in the space below.

Scenario B. Suppose you get inspected and your estimated output is 10. Please calculate what your earnings would be based on your choice of reported output and write this in the space below.

Scenario C. Suppose you get inspected and your estimated output is 30. Please calculate what would your earnings would be based on your choice of reported output and write this in the space below.

*Please raise your hand when you are ready to have your calculations checked or if you have a question.*

### Chance of Inspection Table

---

If the <i>average</i> reported output of the other players in your group is _____ <i>your</i> reported output,	your chance of inspection is . . .
4 or more units <i>less than</i>	0%
3 units <i>less than</i>	5%
2 units <i>less than</i>	10%
1 unit <i>less than</i>	15%
equal to	20%
1 unit <i>greater than</i>	25%
2 units <i>greater than</i>	30%
3 units <i>greater than</i>	35%
4 units <i>greater than</i>	40%
5 units <i>greater than</i>	45%
6 units <i>greater than</i>	50%
7 units <i>greater than</i>	55%
8 units <i>greater than</i>	60%
9 units <i>greater than</i>	65%
10 units <i>greater than</i>	70%
11 units <i>greater than</i>	75%
12 units <i>greater than</i>	80%
13 units <i>greater than</i>	85%
14 units <i>greater than</i>	90%
15 units <i>greater than</i>	95%
16 or more units <i>greater than</i>	100%

---



## References

- Alm, J., et al. (1993). Tax Compliance with Endogenous Audit Selection Rules. 46: 27-45.
- Alm, J., et al. (1992). Why do people pay taxes? 48.
- Alm, J. and M. McKee (1998). Extending the Lessons of Laboratory Experiments on Tax Compliance to Managerial and Decision Economics. 19.
- Alm, J. and M. McKee (2004). Tax compliance as a coordination game. 54: 297-312.
- Bertrand, M., et al. (2004). How Much Should We Trust Differences-in-Differences Estimates? 119.
- Bull, C., et al. (1987). Tournaments and piece rates An experimental study. 95.
- Cason, T. and L. Gangadharan (2006). An experimental study of compliance and leverage in auditing and regulatory enforcement. 44: 352-366.
- Clark, J., et al. (2004). The Good, the Bad, and the Regulator: An Experimental Test of Two Conditional Audit Schemes. 42: 69-87.
- Dechenaux, E., et al. (2006). Caps on bidding in all-pay auctions: comments on the experiments of A. Rapoport and W. Amaldoss. 61: 276-283.
- Evans, M. F., et al. (2009). Regulation with Direct Benefits of Information Disclosure and Imperfect Monitoring. 57: 284-292.
- Fischbacher, U. (2007). z-Tree Zurich Toolbox for Ready-made Economic Experiments. 10: 171-178.
- Friesen, L. (2003). Targeting Enforcement to Improve Compliance with Environmental Regulations. 46: 72-86.
- Greenberg, J. (1984). Avoiding tax avoidance: a (repeated) game-theoretic approach. 32.
- Harford, J. D. (1991). Measurement error and state-dependent pollution control enforcement. 21: 67-81.
- Harrington, W. (1988). Enforcement leverage when penalties are restricted. 37.
- Kezdi, G. (2004). Robust Standard Error Estimation in Fixed-Effect Panel Models. 9: 95-116.
- Knoeber, C. R. and W. N. Thurman (1994). Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production. 12.
- Landsberger, M. and I. Meilijson (1982). Incentive generating state dependent penalty system: the case of income tax evasion. 19.
- Lazear, E. P. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. 89.

Nalbantian, H. R. and A. Schotter (1997). Productivity under group incentives: An experimental study. 87.

Nalebuff, B. J. and J. E. Stiglitz (1983). Prizes and incentives: Towards a general theory of compensation and competition. 14.

Newey, W. K. and K. D. West (1987). A simple, positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. 55.

O'Keeffe, M., et al. (1984). Economic contests: Comparative reward schemes. 2.

Raymond, M. (1999). Enforcement leverage when penalties are restricted: a reconsideration under asymmetric information. 73.

Rickard, J. A., et al. (1982). A tax evasion model with allowance for retroactive penalties. 58.

Stafford, S. L. (2008). Self-Policing in a Targeted Enforcement Regime. 74: 934-951.

van Dijk, F., et al. (2001). Incentive systems in a real effort experiment. 45.

Vossler, C. A. (2008). Analyzing Repeated-Game Economics Experiments: Robust Standard Errors for Panel Data with Serial Correlation. Handbook on Experimental Economics and the Environment B2 - Handbook on Experimental Economics and the Environment. Northampton , MA, Edward Elgar.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. 40: 617-636.